Protein Labeling

DOI: 10.1002/anie.201001240

Enzymatic Site-Specific Functionalization of Protein Methyltransferase Substrates with Alkynes for Click Labeling**

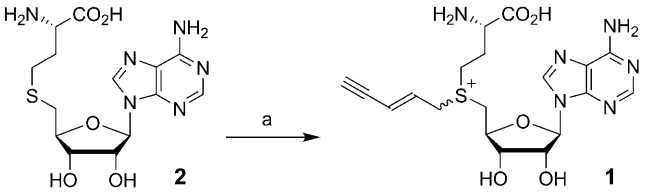
*Wibke Peters, Sophie Willnow, Mike Duisken, Henning Kleine, Thomas Macherey, Kelly E. Duncan, David W. Litchfield, Bernhard Lüscher, and Elmar Weinhold**

Posttranslational modifications of proteins are key to essentially all regulatory processes in cells. Many different modifications, including methylation, have been described for core histones, the protein components of nucleosomes. The modifications occur preferentially on the N-terminal tails and are thought to control the interaction with proteins associated with the regulation of chromatin structure and gene transcription.[1–4] Recent studies have demonstrated that methylation of the side chains of lysine and arginine residues of core histones are associated with specific functional states of promoters. For example, methylation of histone H3 at lysine 9 (H3K9) is a negative mark for gene transcription,[1,2] and trimethylation of histone H3 at lysine 4 (H3K4) is a marker for transcribed promoters.[3,5] Methylation at H3K4 is interconnected with other histone modifications, including dimethylation of histone H3 at arginine 2 (H3R2), a transcriptionally negative mark which inhibits methylation of H3K4.[6,7]

Protein methyltransferases (MTases) transfer the activated methyl group from the cofactor *S*-adenosyl-L-methionine (AdoMet or SAM) mainly to lysine and arginine side chains in their protein substrates. These enzymes are often sequence-specific; for example, mixed-lineage leukaemia (MLL) histone MTase complexes trimethylate H3K4.[8] Methylation of lysine residues is a dynamic, reversible modification involving MTases and demethylases.[9] The main protein MTase substrates described are core histones and a few proteins associated with gene transcription.[10] Comprehensive analyses of MTase substrates are lacking, at least in part because the methyl group is a poor reporter. Antibodies seem to recognize methylated amino acids only in a context-dependent manner, that is, in combination with the underlying peptide sequence. Therefore we thought to develop alternative methods to identify MTase substrates.

Recently, we reported on synthetic double-activated AdoMet analogues with allylic and propargylic methyl group replacements for site-specific DNA modification by DNA MTases.[11,12] Such analogues also function as cofactors for small molecule MTases.[13] Compared to aziridinium-based AdoMet analogues,[14] these cofactors have the advantage that strong product inhibitors are not formed during the MTase-catalyzed reaction. When an amino function was appended to the propargylic side chain, it was possible to couple *N*-hydroxysuccinimde (NHS)-activated reporters to the modified DNA in a second step.[15]

Since introduction of amino groups is generally not productive for the analysis of proteins, we designed the new AdoMet-based cofactor AdoEnYn (**1**; Scheme 1), in which the methyl group of AdoMet is replaced by a pent-2-en-4-ynyl side chain. The double bond in β position to the sulfonium center compensates unfavorable steric effects within the S_N2-like transition state by conjugative stabilization[11] and the terminal alkyne serves as a reactive bio-orthogonal functionality for chemical protein manipulations. This design was necessary because a corresponding AdoMet analogue with a prop-2-ynyl group proved unstable especially under basic conditions. AdoEnYn (**1**) was synthesized in a one-step reaction from *S*-adenosyl-L-homocysteine (AdoHcy, **2**) and the corresponding activated alcohol (see the Supporting Information for experimental details). Protecting groups are not needed in this nucleophilic substitution reaction because



Scheme 1. Synthesis of AdoEnYn (**1**). Conditions: a) (*E*)-Pent-2-en-4-ynyl methanesulfonate, HCO_2H, CH_3CO_2H, RT, 14 h, 50%.

[*] Dr. W. Peters, S. Willnow, T. Macherey, Prof. Dr. E. Weinhold
Institute of Organic Chemistry, RWTH Aachen University
Landoltweg 1, 52056 Aachen (Germany)
Fax: (+49) 241-80-92528
E-mail: elmar.weinhold@oc.rwth-aachen.de
Homepage: http://www.weinholdgroup.rwth-aachen.de
Dr. W. Peters, Dr. H. Kleine, Prof. Dr. B. Lüscher
Institute of Biochemistry and Molecular Biology
Medical School, RWTH Aachen University
Pauwelsstrasse 30, 52074 Aachen (Germany)
Dr. M. Duisken
Leco Instruments GmbH, Mönchengladbach (Germany)
Dr. K. E. Duncan, Prof. Dr. D. W. Litchfield
Department of Biochemistry
Schulich School of Medicine and Dentistry
University of Western Ontario, London ON (Canada)

[**] We thank Xiaodong Cheng for providing the Dim-5 expression plasmid and Kerstin Glensk for technical assistance in preparing the AdoEnYn cofactor analogue. This work was supported by the RWTH Aachen University (graduate stipend to W.P.), the Canadian Institutes of Health Research (grant MOP 37854 to D.W.L.), the Deutsche Forschungsgemeinschaft (grant LU 466/13-1 to B.L. and grants WE 1453/4-1 and 4-2 to E.W.), and the Excellence Initiative of the German Federal and State Governments.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201001240.

the acidic conditions lead to protonation and hence transient protection of all other nucleophilic positions in AdoHcy (**2**) except the sulfur atom.

To test whether the novel cofactor AdoEnYn (**1**) can be used by protein MTases, we chose the SET-domain-containing enzyme Dim-5 from *Neurospora crassa*, which specifically trimethylates H3K9.[16] We used peptide **3** as a substrate which comprises amino acids 1–15 of histone H3 and an N-terminally coupled fluorescein residue for detection (Figure 1a). Peptide **4** modified with the unpolar pent-2-en-4-ynyl side chain showed an increased retention time in the reverse-phase HPLC analysis (Figure 1b). Product analysis by high-resolution LC–ESI TOF and MALDI-TOF/TOF mass spectrometry (MS) confirmed that peptide **4** was modified with one pent-2-en-4-ynyl side chain (see Figure S4 in the Supporting Information) site-specifically at lysine 9 (Figure 1c).

Transfer of the pent-2-en-4-ynyl side chain introduces a terminal alkyne into the peptide substrate, which can be used for further modifications using bio-orthogonal CuAAC click chemistry (CuAAC: Cu^{I}-catalyzed azide–alkyne 1,3-dipolar cycloaddition).[17] The reaction between alkyne-modified peptide **4** and azide-derivatized biotin **5** (Figure 2a) was analyzed by reverse-phase HPLC and MS (Figure 2b). The biotinylated peptide **6** showed an increased retention time and the mass determined by high-resolution LC–ESI TOF MS was in good agreement with the predicted structure (see Figure S5 in the Supporting Information).

To further evaluate the system, we used full-length histone H3 for the two-step biotinylation procedure. Histone H3 was modified by Dim-5 using AdoEnYn (**1**) and subsequently labeled with the biotin azide **5** prior to detection with peroxidase-conjugated avidin after SDS-PAGE and western blotting. A specific signal corresponding to the labeled

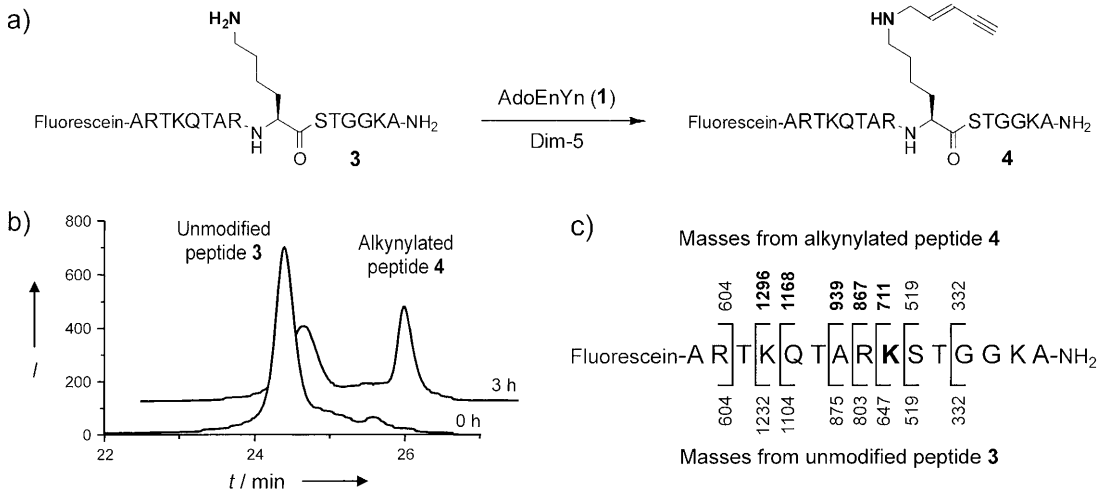


Figure 1. Enzymatic transfer of the pent-2-en-4-ynyl side chain from the AdoMet-based cofactor AdoEnYn (**1**) to a model peptide. a) Modification of a histone H3 peptide (amino acids 1–15) containing an N-terminal fluorescein group **3,** the cofactor analogue AdoEnYn (**1**), and the histone MTase Dim-5 at lysine 9. b) Analysis of the enzymatic reaction by reverse-phase HPLC. Peptides were observed by fluorescence detection; fluorescence intensities (*I*) were measured at $\lambda^{Ex} = 440$ nm and $\lambda^{Em} = 510$ nm. Conversion of the reaction was about 55%; full HPLC traces are given in Figure S3 in the Supporting Information. c) MALDI-TOF/TOF MS sequencing results of modified peptide **4** (upper masses) and unmodified peptide **3** (lower masses). Brackets indicate observed fragments; corresponding fragments containing lysine 9 show a mass difference of 64 in agreement with the pent-2-en-4-ynyl modification.

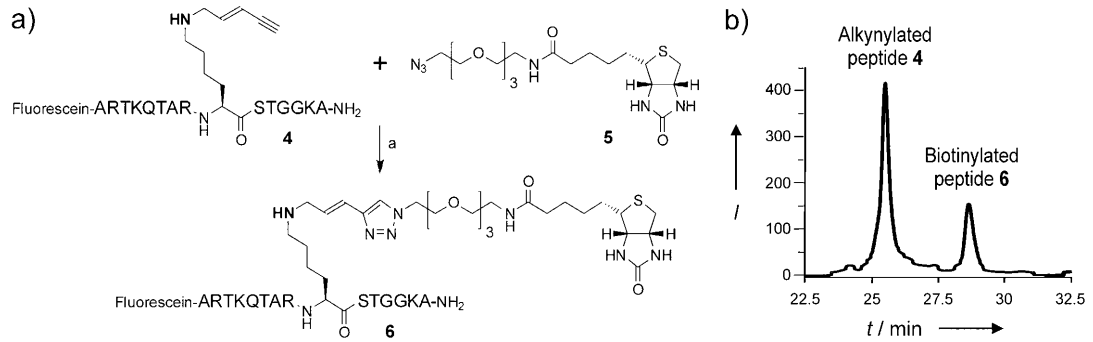


Figure 2. Labeling of alkynylated histone H3 peptide **4** with biotin azide **5** by CuAAC click chemistry. a) Synthesis of the biotinylated peptide **6**. Conditions: a) $CuSO_4$, sodium ascorbate, L-proline, 30°C, overnight. b) Analysis of the chemical modification reaction by reverse-phase HPLC. Peptides were observed by fluorescence detection; fluorescence intensities (*I*) were measured at $\lambda^{Ex} = 440$ nm and $\lambda^{Em} = 510$ nm. Conversion of the reaction was about 30%; full HPLC traces are given in Figure S3 in the Supporting Information.

histone H3 was detected, while negligible signals, probably arising from a slow non-enzymatic transfer, were seen in the controls (Figure 3a). Dim-5 showed automodification, which

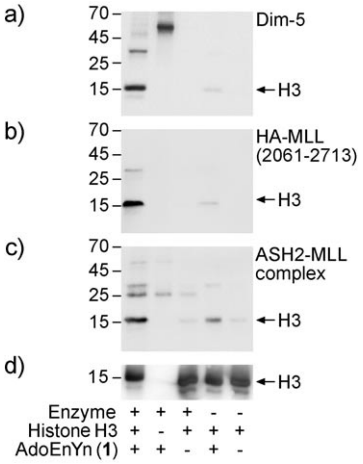


Figure 3. Histone H3 was enzymatically modified with AdoEnYn (**1**) by using with a) Dim-5, b) immunoprecipitated HA-MLL (2061–2713) expressed in HEK293 cells, or c) a tandem-affinity-purified human ASH2-MLL complex from HEK293 cells. The modified proteins were then coupled with azide-derivatized biotin **5** prior to their analysis by SDS-PAGE and western blotting using peroxidase-conjugated avidin. d) Histone H3 was detected with a specific antibody for control.

was strongly reduced in the presence of histone H3, suggesting that the latter is a better substrate than Dim-5 itself. Taken together, these results demonstrate that AdoEnYn (**1**) serves as a cofactor for Dim-5 and allows sequence-specific modification of lysine 9 of both a histone H3 N-terminal peptide and full-length histone H3. Finally, the transferred terminal alkyne can be modified by CuAAC click chemistry, which allows specific detection of the MTase substrate.

Dim-5 is a fungal, monomeric protein MTase, whereas human protein MTases are frequently complexed with additional subunits critical for enzymatic activity. MLL MTase complexes are of considerable interest because some subunits are deregulated in cancer, including MLLs, the tumor suppressor menin, and the oncoprotein ASH2.[8,18] Therefore, we tested MLL MTases with AdoEnYn (**1**). An HA-tagged C-terminal fragment of MLL4 (amino acids 2061–2713 containing the catalytic SET domain) was expressed in HEK293T cells, assuming that it would complex with endogenous subunits.[19] Indeed, this complex was active when tested with [*methyl*-^{3}H]AdoMet (data not shown) and capable of using AdoEnYn (**1**) to modify histone H3 (Figure 3b). In addition, a purified ASH2-MLL complex that trimethylates lysine 4 of histone H3[6] was able to transfer the pent-2-en-4-ynyl side chain from the synthetic cofactor **1** to histone H3 (Figure 3c). Thus, different protein MTases are capable of utilizing AdoEnYn (**1**) as a cofactor. This allows the specific transfer of a chemically modifiable alkyne group to substrate proteins, which then can be labeled by CuAAC click chemistry using azide-derivatized reporter groups under mild conditions.

In conclusion, our findings demonstrate that alternative cofactors for protein MTases can be designed and synthesized. It is remarkable that AdoEnYn (**1**) is suitable for the mammalian ASH2-MLL complex, which is composed of multiple subunits. Since several subunits of this complex are involved in tumorigenesis,[8,18] an important question is whether the oncogenic properties of the ASH2-MLL complex are mediated by its ability to trimethylate H3K4, or whether additional substrates exist that are critical for tumor development. The lack of efficient methods has thus far prevented a comprehensive analysis of ASH2-MLL substrates. This is also true for other protein MTases. Recent studies using peptide arrays suggest that consensus sequences for MTase substrates can be derived.[20] Such consensus sequences permit computer-supported identification of potential substrates but require considerable verification experiments. The methods and the compounds defined here will now allow the development of screening procedures to identify novel MTase substrates. Furthermore, we envisage that when used in conjunction with protein arrays, our novel approach will enable comprehensive studies to define the methylated proteome.

Received: March 1, 2010
Published online: June 22, 2010

Keywords: click chemistry · cofactors · epigenetics · proteins · transferases

[1] S. L. Berger, *Nature* **2007**, *447*, 407 – 412.
[2] T. Kouzarides, *Cell* **2007**, *128*, 693 – 705.
[3] A. Shilatifard, *Curr. Opin. Cell Biol.* **2008**, *20*, 341 – 348.
[4] R. J. Sims, D. Reinberg, *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 815 – 820.
[5] a) T. A. Milne, S. D. Briggs, H. W. Brock, M. E. Martin, D. Gibbs, C. D. Allis, J. L. Hess, *Mol. Cell* **2002**, *10*, 1107 – 1117; b) R. Schneider, A. J. Bannister, F. A. Myers, A. W. Thorne, C. Crane-Robinson, T. Kouzarides, *Nat. Cell Biol.* **2004**, *6*, 73 – 77; c) A. Yokoyama, Z. Wang, J. Wysocka, M. Sanyal, D. J. Aufiero, I. Kitabayashi, W. Herr, M. L. Cleary, *Mol. Cell. Biol.* **2004**, *24*, 5639 – 5649.
[6] E. Guccione, C. Bassi, F. Casadio, F. Martinato, M. Cesaroni, H. Schuchlautz, B. Lüscher, B. Amati, *Nature* **2007**, *449*, 933 – 937.
[7] a) D. Hyllus, C. Stein, K. Schnabel, E. Schiltz, A. Imhof, Y. Dou, J. Hsieh, U.-M. Bauer, *Genes Dev.* **2007**, *21*, 3369 – 3380; b) A. Kirmizis, H. Santos-Rosa, C. J. Penkett, M. A. Singer, M. Vermeulen, M. Mann, J. Bähler, R. D. Green, T. Kouzarides, *Nature* **2007**, *449*, 928 – 932.
[8] A. V. Krivtsov, S. A. Armstrong, *Nat. Rev. Cancer* **2007**, *7*, 823 – 833.
[9] R. J. Klose, Y. Zhang, *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 307 – 318.
[10] J. Huang, S. L Berger, *Curr. Opin. Genet. Dev.* **2008**, *18*, 152 – 158.
[11] C. Dalhoff, G. Lukinavicius, S. Klimasauskas, E. Weinhold, *Nat. Chem. Biol.* **2006**, *2*, 31 – 32.
[12] C. Dalhoff, G. Lukinavicius, S. Klimasauskas, E. Weinhold, *Nat. Protoc.* **2006**, *1*, 1879 – 1886.
[13] a) H. Stecher, M. Tengg, B. J. Ueberbacher, P. Remler, H. Schwab, H. Griengl, M. Gruber-Khadjawi, *Angew. Chem.* **2009**, *121*, 9710 – 9712; *Angew. Chem. Int. Ed.* **2009**, *48*, 9546 – 9548; b) B. W. K. Lee, H. G. Sun, T. Zang, B. J. Kim, J. F. Alfaro, Z. S. Zhou, *J. Am. Chem. Soc.* **2010**, *132*, 3642 – 3643.

[14] T. Osborne, R. L. Weller Roska, S. R. Rajski, P. R. Thompson, *J. Am. Chem. Soc.* **2008**, *130*, 4574 – 4575.
[15] G. Lukinavičius, V. Lapiene, Z. Stasevskij, C. Dalhoff, E. Weinhold, S. Klimasauskas, *J. Am. Chem. Soc.* **2007**, *129*, 2758 – 2759.
[16] X. Zhang, H. Tamaru, S. I. Khan, J. R. Horton, L. J. Keefe, E. U. Selker, X. Cheng, *Cell* **2002**, *111*, 117 – 127.
[17] a) V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem.* **2002**, *114*, 2708 – 2711; *Angew. Chem. Int. Ed.* **2002**, *41*, 2596 – 2599; b) C. W. Tornøe, C. Christensen, M. Meldal, *J. Org. Chem.* **2002**, *67*, 3057 – 3064; c) R. Breinbauer, M. Köhn, *ChemBioChem* **2003**, *4*, 1147 – 1149.
[18] a) A. Yokoyama, T. C. P. Somervaille, K. S. Smith, O. Rozenblatt-Rosen, M. Meyerson, M. L. Cleary, *Cell* **2005**, *123*, 207 – 218; b) J. Lüscher-Firzlaff, I. Gawlista, J. Vervoorts, K. Kapelle, T. Braunschweig, G. Walsemann, C. Rodgarkia-Schamberger, H. Schuchlautz, S. Dreschers, E. Kremmer, R. Lilischkis, C. Cerni, A. Wellmann, B. Lüscher, *Cancer Res.* **2008**, *68*, 749 – 758.
[19] M. M. Steward, J.-S. Lee, A. O'Donovan, M. Wyatt, B. E. Bernstein, A. Shilatifard, *Nat. Struct. Mol. Biol.* **2006**, *13*, 852 – 854.
[20] a) P. Rathert, A. Dhayalan, M. Murakami, X. Zhang, R. Tamas, R. Jurkowska, Y. Komatsu, Y. Shinkai, X. Cheng, A. Jeltsch, *Nat. Chem. Biol.* **2008**, *4*, 344 – 346; b) P. Rathert, X. Zhang, C. Freund, X. Cheng, A. Jeltsch, *Chem. Biol.* **2008**, *15*, 5 – 11.

